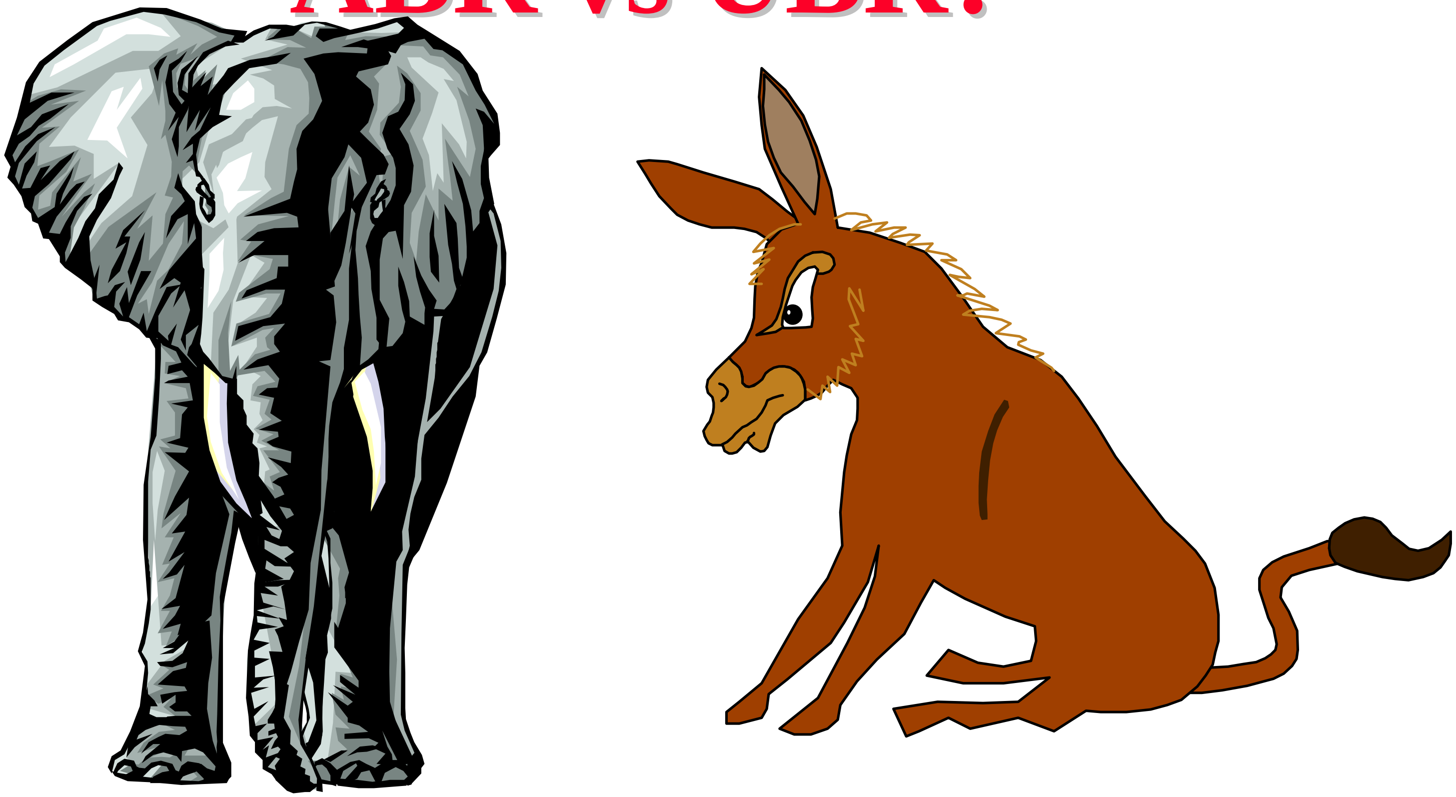
Design Issues in Traffic Management for the ATM UBR+ Service for TCP over Satellite Networks: Final Report

Raj Jain

Raj Jain is now at
Washington University in Saint Louis
Jain@cse.wustl.edu
http://www.cse.wustl.edu/~jain/

- Statement of Work: TCP over UBR Issues to Study
- TCP Policies using Bursty Traffic
  - WWW Model
  - Full Factorial Experimental Design and Analysis
- Summary of Other Results

Internet Protocols over ABR vs UBR?



- Intelligent transport or not?

Goals: Issues

1. Analyze Standard Switch and End-system Policies
2. Design Switch Drop Policies
3. Quantify Buffer Requirements in Switches
4. UBR with VBR Background
5. Performance of Bursty Sources
6. Changes to TCP Congestion Control
7. Optimizing the Performance of SACK TCP

Non-Goals

- Does not cover non-UBR issues.
- Does not cover ABR issues.
- Does not include non-TM issues.

Status

1. Analyze Standard Switch and End-system Policies[1]
2. Design Switch Drop Policies[2]
3. Quantify Buffer Requirements in Switches[1]
4. UBR with VBR Background: GFR[2] and GR[1]
5. Performance of Bursty Sources[3]
6. Changes to TCP Congestion Control[2]
7. Optimizing the Performance of SACK TCP[2]

Status: [1]=Presented at the 1st meeting,

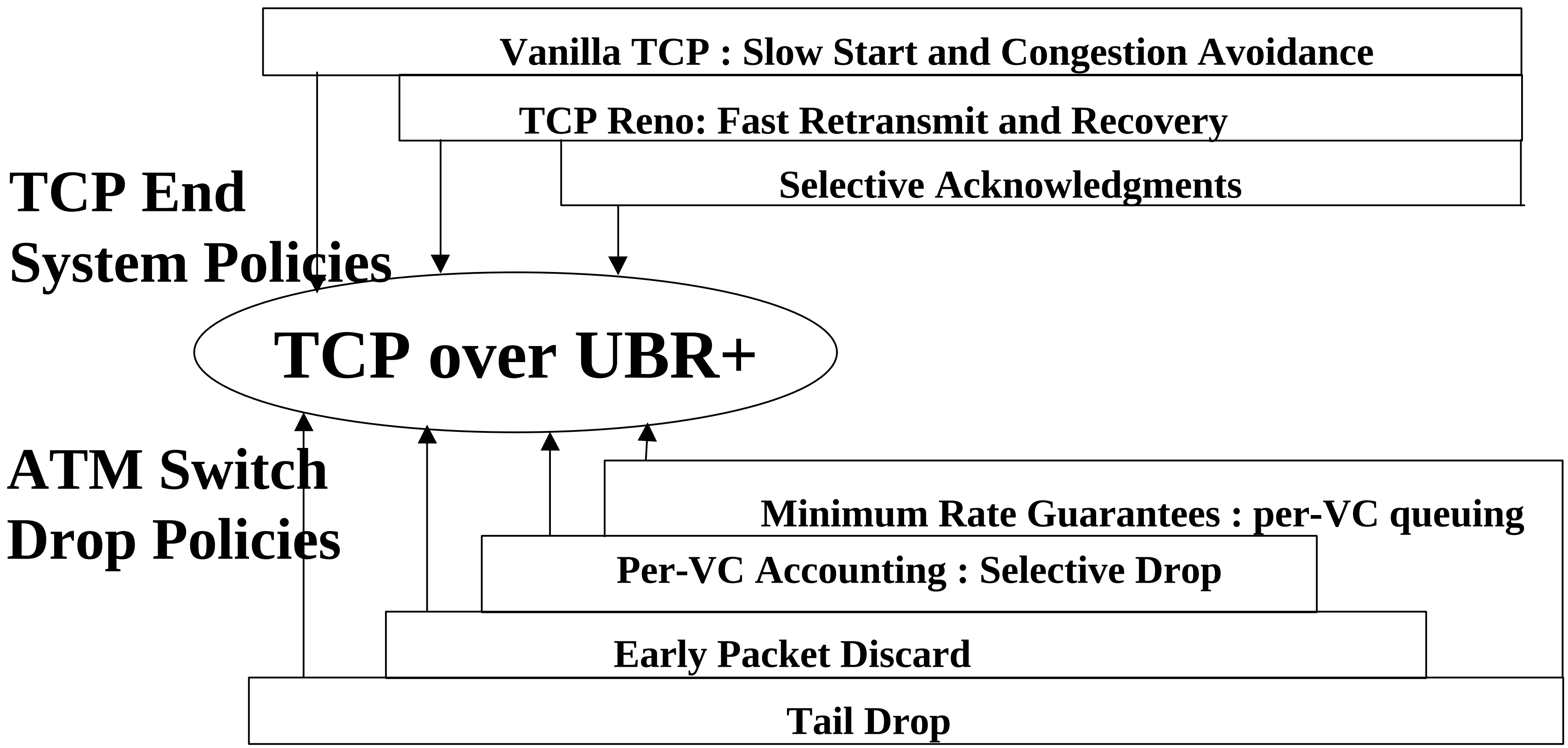
[2]=2nd Meeting [3]=Presenting now

Task 5: Goals

- Analyze the effect of three factors on WAN, MEO, GEO:

1. TCP Flavors
   - Vanilla: Slow start and congestion avoidance
   - Fast retransmit and recovery (Reno)
   - New Reno
   - Selective Acknowledgements
2. Switch Drop Policies
   - EPD
   - Per-VC accounting
3. Buffer Size: 0.5, 1, 2 × RTT-bandwidth product

TCP over UBR+



Vanilla TCP : Slow Start and Congestion Avoidance
TCP Reno: Fast Retransmit and Recovery
Selective Acknowledgments
TCP End System Policies
TCP over UBR+
ATM Switch Drop Policies
Minimum Rate Guarantees : per-VC queuing
Per-VC Accounting : Selective Drop
Early Packet Discard
Tail Drop

Policies

LEO, MEO, GEO
# of TCP Sources
Buffer Sizes

End-System Policies

Switch Policies

		No FRR	FRR	New Reno	SACK + New Reno
No EPD					
	Plain EPD				
EPD	Selective Drop				
	Fair Buffer Allocation				

SPECWeb 96 WWW Model

- Majority of traffic on the Internet is WWW
- Developed by Standard Performance Evaluation Corporation (SPEC), a consortium similar to the ATM Forum for performance benchmarking
- SPECMark, SPEC CPU95, SPECInt95, SPEC SFS
- SPECWeb96 is for benchmarking WWW servers
- Ref: http://www.specbench.org/ost/web96/webpaper.html

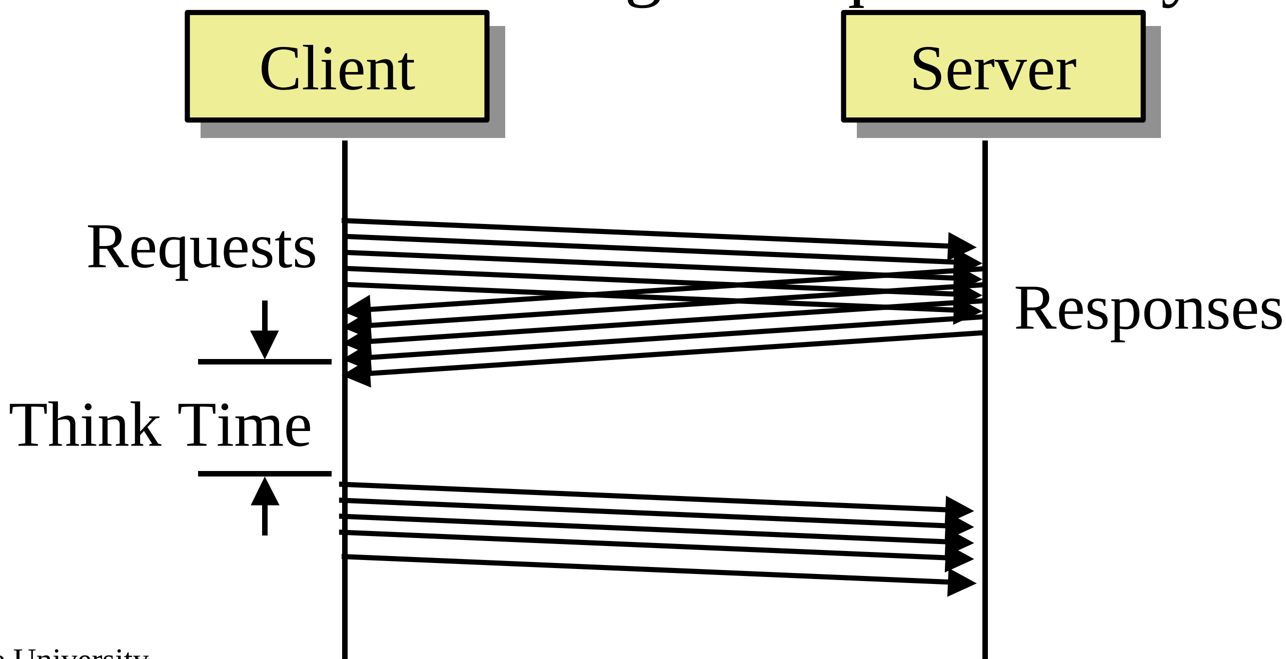
Modified SPECWeb96

Class 0	Class 1	Class 2	Class 3	Class 4
p = 0.2	p = 0.28	p = 0.40	p = 0.112	p = 0.008
0.1 kB 0.2 kB ... 0.9 kB	1 kB 2 kB ... 9 kB	10 kB 20 kB ... 90 kB	100 kb 200 kB ... 900 kB	1 MB 2 MB ... 9 MB

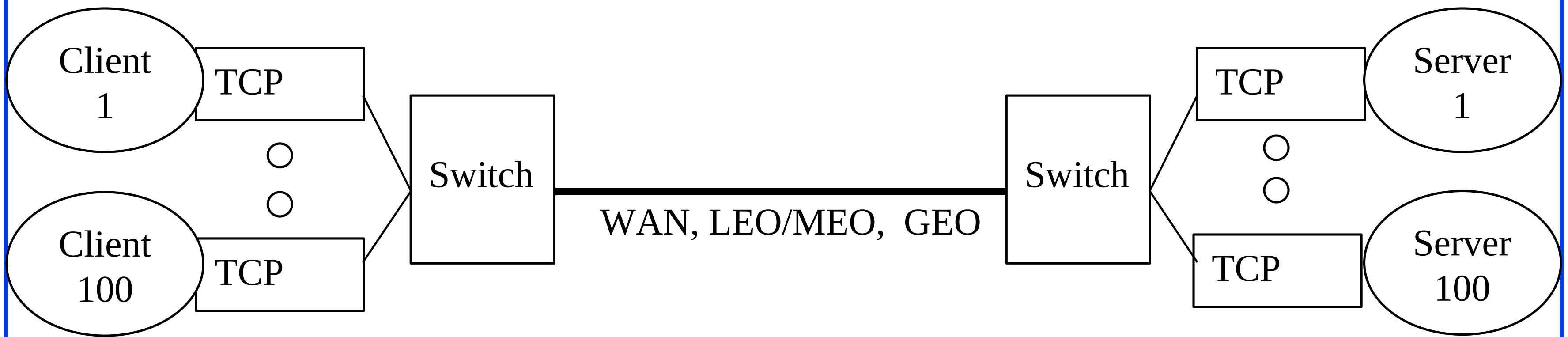
- Each web page consists of one index page and 4 images.
- First column: Index page (p = 1/5)
- Other columns: p = 0.8

Modified SPECWeb 96

- Average file size = 120.3 KB
- Bandwidth per client = 0.48 Mbps
- HTTP 1.1 ⇒ All components of a web page are fetched in one TCP connection.
- A client makes on average 5 requests every 10s.



N Client-Server Configuration



- 1 client per server, N clients and servers, N=100
- RTTs for WAN,multiple-hop LEO/Single-hop MEO and GEO link: 10ms, 200ms and 550ms
- Inter-switch link Bandwidth: 45 Mbps (T3)
- Simulation Time = 100secs i.e. 10 cycles of client requests

TCP Parameters

- MSS = 1024 (WAN), 9180 (LEO/MEO, GEO) bytes
- RCV_WND > RTT × Bandwidth
- "Silly Window Syndrome Avoidance" disabled, since WWW requests must be sent right away.
- Initial SS_THRESH = RTT × Bandwidth [HOE96]
- TCP delay ACK timer is NOT set ⇒ No ack delay
- TCP max window scaled using window scaling option
- TCP timer granularity = 100 ms

Switch Parameters

Link Type (RTT)	RTT-bandwidth product (cells)	Switch Buffer Sizes (cells)
WAN (10 ms)	1062	531, 1062, 2300
Multiple-Hop LEO/Single-Hop MEO (200 ms)	21230	10615, 21230, 42460
Single-Hop GEO (550 ms)	58380	29190, 58380, 116760

Analysis Technique

Factors	Levels
TCP Flavor	Vanilla, Reno, NewReno, SACK
Buffer Size	0.5 RTT, 1 RTT, 2 RTT
Switch Drop Policy	EPD, SD

- Separate analysis for Efficiency and Fairness results.
- $y_{ijk} = \mu + \underbrace{\alpha_i + \beta_j + \chi_k}_{} + \underbrace{\delta_{ij} + \gamma_{jk} + \phi_{ik}}_{} + \varepsilon_{ijk}$

Observation = Mean + Main Effects + Interaction + Error

Analysis Technique (contd.)

- $\Sigma y_i^2 = n\mu^2 + \Sigma\alpha_i^2 + \Sigma\beta_j^2 + \Sigma\chi_k^2 + \Sigma\delta_{ij}^2 + \Sigma\gamma_{jk}^2 + \Sigma\phi_{ik}^2 + \Sigma\varepsilon_i^2$
  $SS_Y = SS_{Mean} + SS_{Main\ Effects} + SS_{Interaction} + SS_{Error}$
- **Overall Mean μ**: Mean of all values
- **Overall Variation**: Sum of squares of Y
- **Main Effects**: Means of a particular level and factor
- **First Order Interactions**: Interactions between 2 levels of any two factors.
- **Allocation of Variations**: % of the overall variation explained by each effect
- **Confidence Intervals of Effects**: Is the main effect statistically significant?

Results: WAN Efficiency

- $\Sigma y_i^2 = n\mu^2 + \Sigma\alpha_i^2 + \Sigma\beta_j^2 + \Sigma\chi_k^2 + \Sigma\delta_{ij}^2 + \Sigma\gamma_{jk}^2 + \Sigma\phi_{ik}^2 + \Sigma\varepsilon_i^2$
  $SS_Y = SS_{Mean} + SS_{Main\ Effects} + SS_{Interaction} + SS_{Error}$
  100% = (57% TCP + 30% Buffers + 0% Drop) + (9% TCP×Buffer + 2% TCP×Drop + 0% Buffer×Drop) + 0.3% error
- TCP flavor is most important factor (57% variation)
  - NewReno and SACK show best performance
  - SACK is worse for low buffer (high congestion)

WAN Efficiency (Cont)

- Buffer size is next important factor (30% of variation)
  - Increase in buffer size increases efficiency
  - More room for improvement for Vanilla and Reno
  - Buffer size of 1 RTT is sufficient. This may be related to the number of TCP connections.
- Drop policies have little effect
  - For small buffer, SD is better than EPD

Results: MEO Efficiency

- TCP flavor explains 57% of variation
  - SACK clearly gives best performance
  - Importance of SACK increases as delay increases
- Buffer size is next important factor (22% of variation)
  - Increase in buffer size increases efficiency
  - More room for improvement for Vanilla and Reno
  - Buffer size of 0.5 RTT is sufficient
- Drop policies have little effect

Results: GEO Efficiency

- TCP flavor explains 70% of variation
  - SACK clearly gives the best performance
- Buffer size is the next important factor (14% of variation)
  - Increase in buffer size increases efficiency
  - More room for improvement for Vanilla and Reno
  - Buffer size of 0.5 RTT is sufficient
- Drop policies have little effect

Overall Results: Efficiency

- End system policies have more effect as delay increases
  - SACK is generally best esp. for long delay
  - NewReno may be better for lower delay and severe congestion
- Drop policies have more effect on lower delays or smaller buffer sizes.
- Buffer size: Larger buffers improve performance. 0.5 RTT to 1 RTT buffers sufficient. More does not help. Optimal buffer size may be related to number of TCPs.

Overall Results: Fairness

- End system policies:
  - SACK hurts fairness for lower delay and smaller buffers (particularly if the buffer sizes are small compared to number of TCP sources)
- Drop policies do not have much effect unless delay is lower and buffers are small.
- Buffer size has more effect on longer delays
  - Increase in buffer size increases fairness. But for sufficiently large buffers, this effect is negligible.

Complete Project Summary

- All tasks successfully completed

Task 1 + 2 (Policies) Results

- In LANs, switch improvements (PPD, EPD, SD, FBA) have more impact than end-system improvements (Slow start, FRR, New Reno, SACK). Different variations of increase/decrease have little impact due to small window sizes.
- In satellite networks, end-system improvements have more impact than switch-based improvements
- FRR hurts in satellite networks.
- Fairness depends upon the switch drop policies and not on end-system policies

Policies (Continued)

- In Satellite networks:
  - SACK helps significantly
  - Switch-based improvements have relatively less impact than end-system improvements
  - Fairness is not affected by SACK
- In LANs:
  - Previously retransmitted holes may have to be retransmitted on a timeout
    $\Rightarrow$ SACK can hurt under extreme congestion.

3. Buffer Requirements: Results

- Very small buffer sizes result in low efficiency
- Moderate buffer sizes (less than 1 RTT)
  - Efficiency increases with increase in buffer size
  - Efficiency asymptotically approaches 100%
- 0.5*RTT buffers provide sufficiently high efficiency (98% or higher) for SACK TCP over UBR even for a large number of TCP sources

Task 4a. Guaranteed Frame Rate (GFR)

- UBR with minimum cell rate (MCR)
  ⇒ UBR+
- Frame based service
  - Complete frames are accepted or discarded in the switch
  - Traffic shaping is frame based.
    All cells of the frame have CLP =0 or CLP =1
  - All frames below MCR are given CLP =0 service.
    All frames above MCR are given best effort (CLP =1) service.

Task 4b. Guaranteed Rate

- Guaranteed Rate (GR): Reserve a small fraction of bandwidth for UBR class.

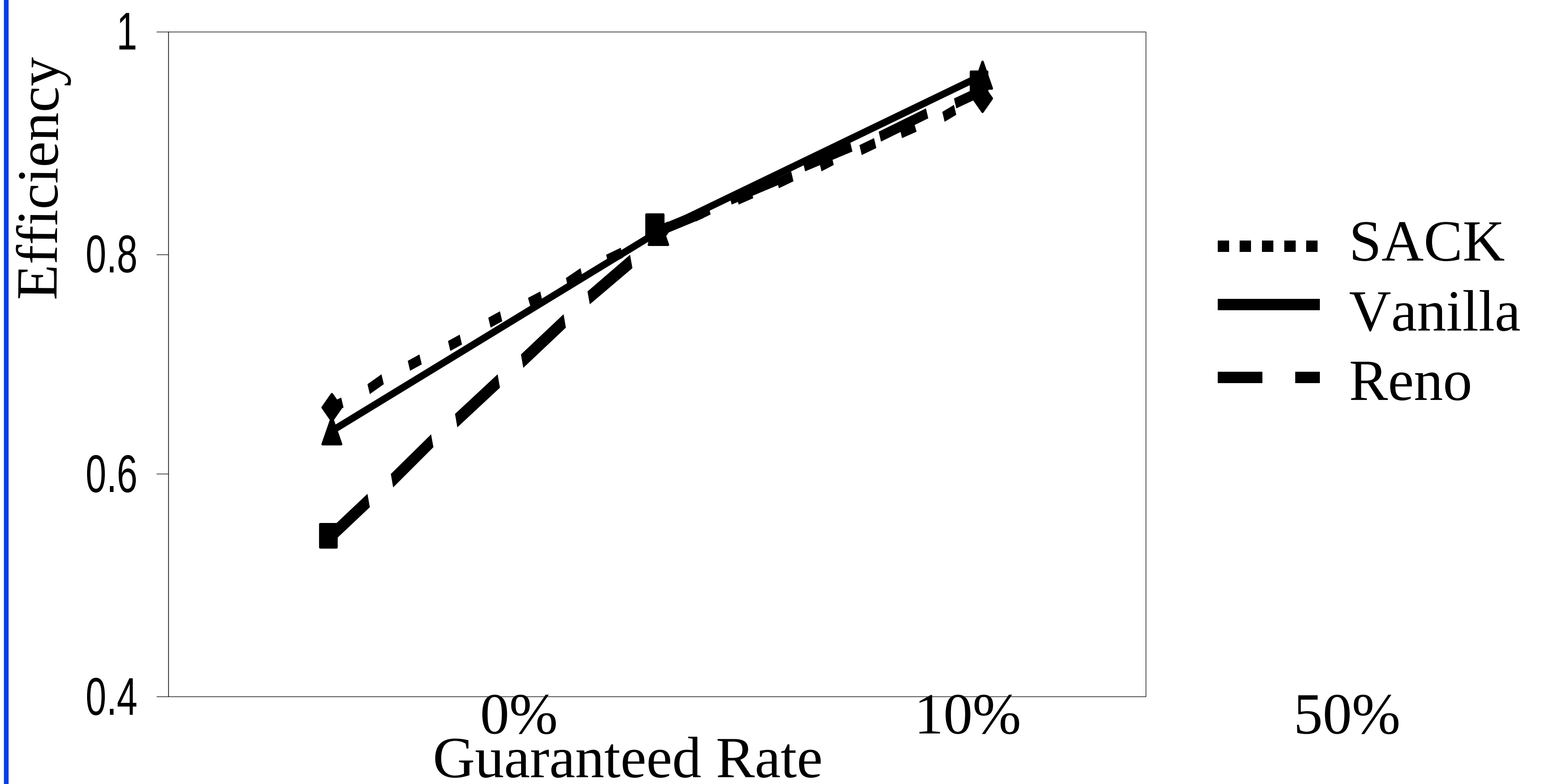
GR	**GFR**
per-class reservation	per-VC reservation
per-class scheduling	per-VC accounting/scheduling
No new signaling	Need new signaling
Can be done now	In TM4+

4b. Guaranteed Rate: Results

- Guaranteed rate is helpful in WANs.
- For WANs, the effect of reserving 10% bandwidth for UBR is more than that obtained by EPD, SD, or FBA
- For LANs, guaranteed rate is not so helpful. Drop policies are more important.
- For Satellites, end-system policies seem more important.

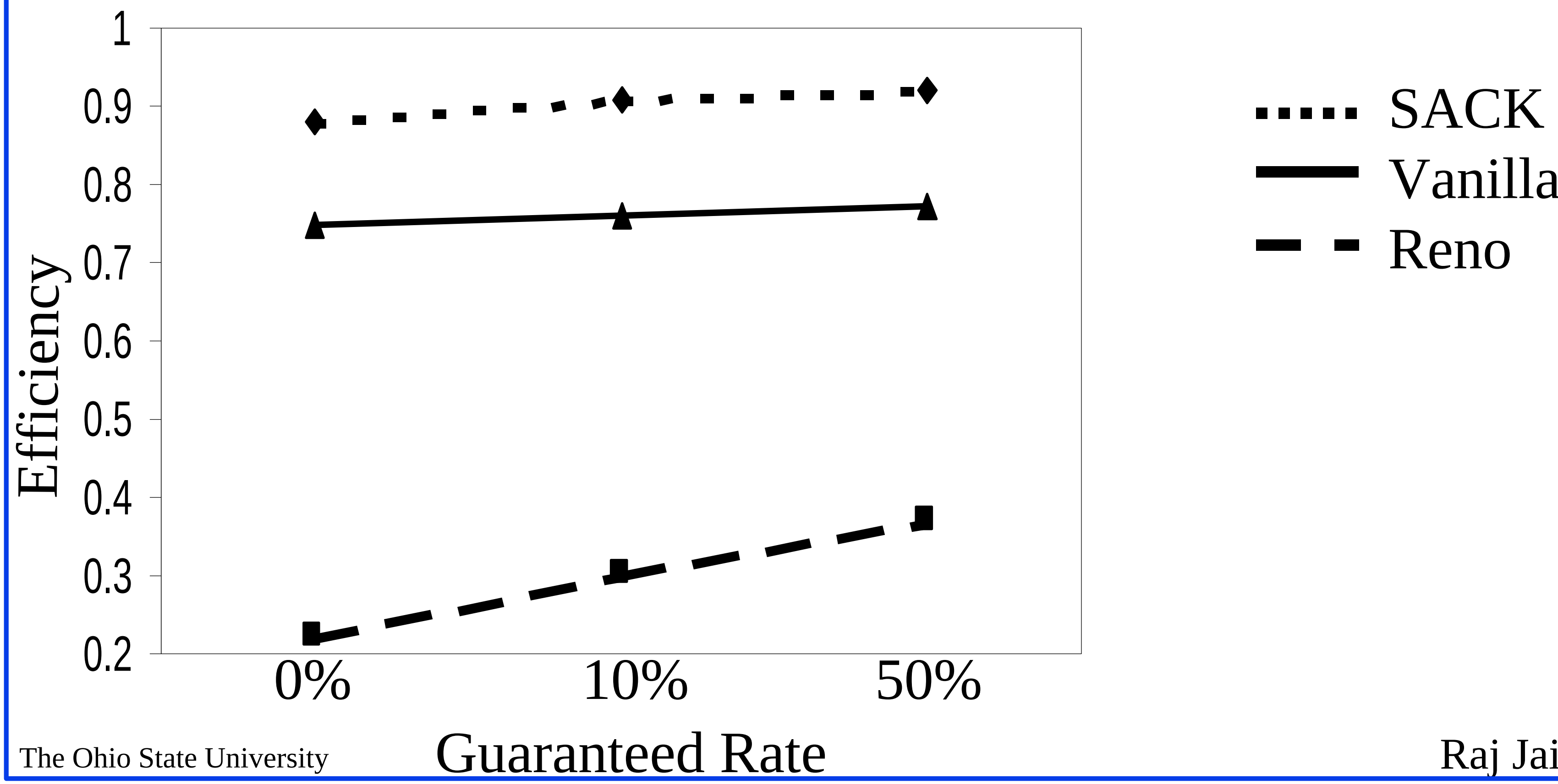
WANs: TCP Mechanisms vs GR

- 5 and 15 Sources, 1 and 3 RTT buffers, SD and EPD, all averaged together



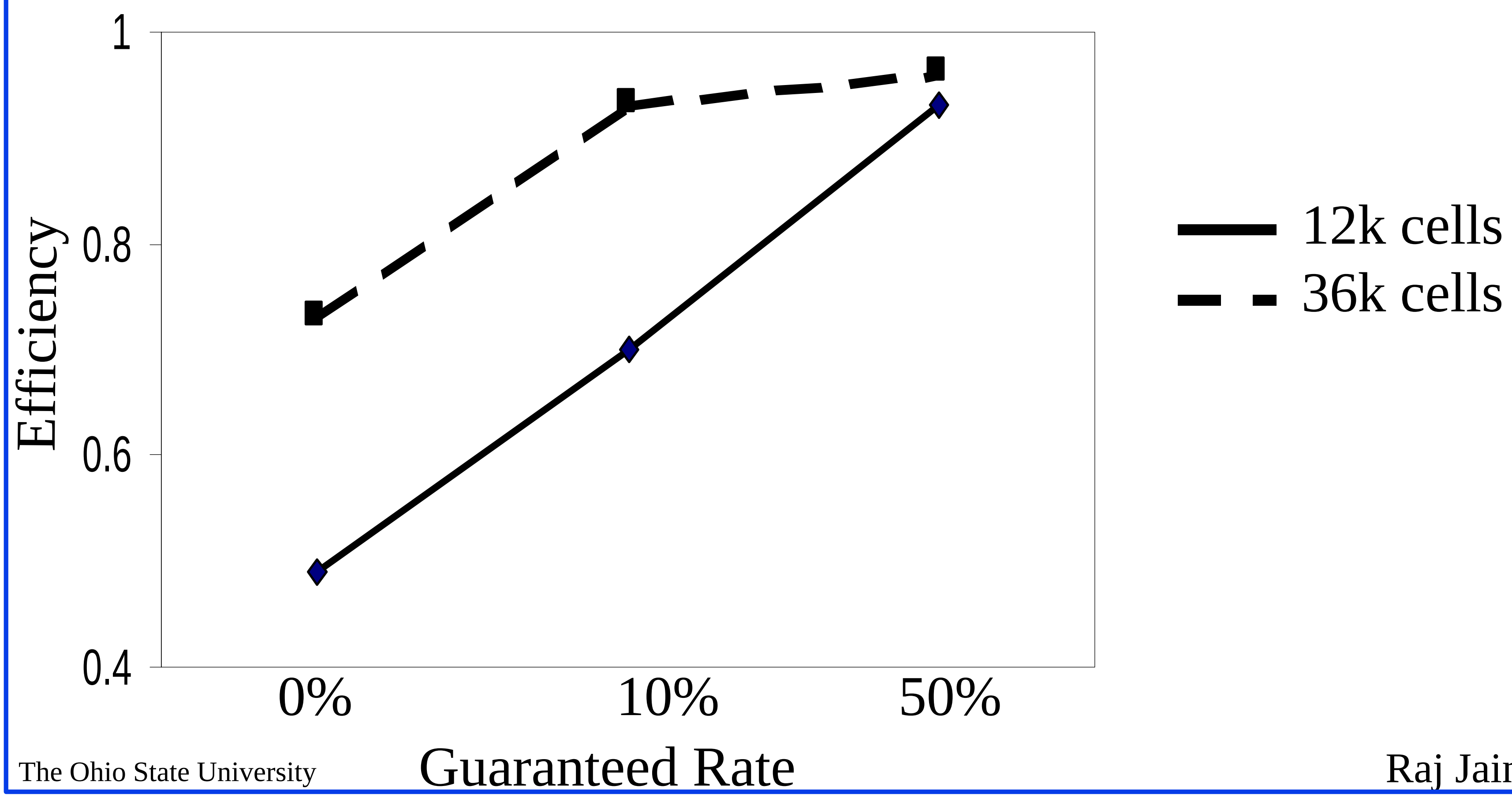
GEO: TCP Mechanisms vs GR

- 5 sources, 1 and 3 RTT buffers, and SD and EPD
- Conclusion: Effect of GR is not as significant as in WANs



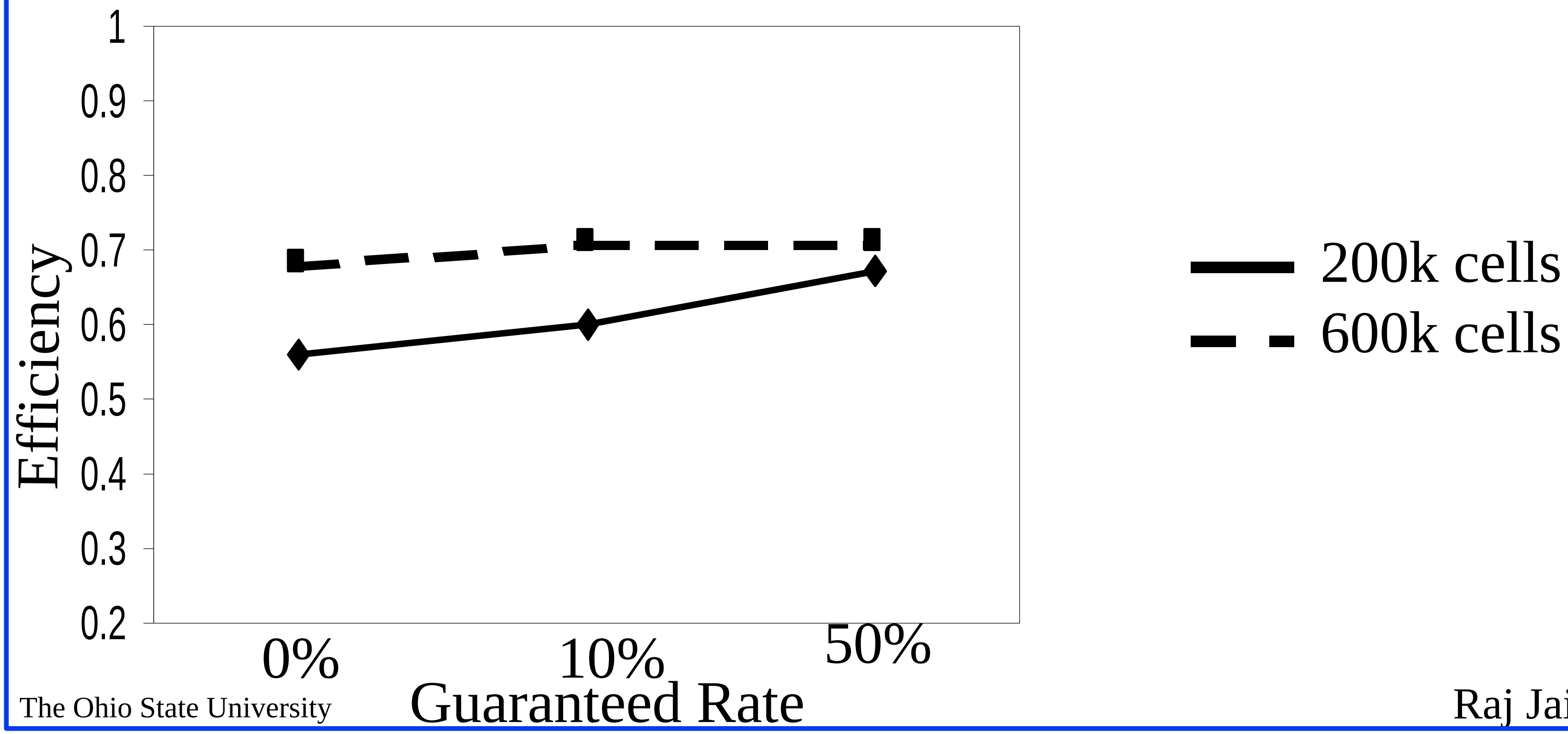
WANs: Buffers vs GR

- 5 and 15 sources, Vanilla, SACK, FRR
  SD and EPD, all average together

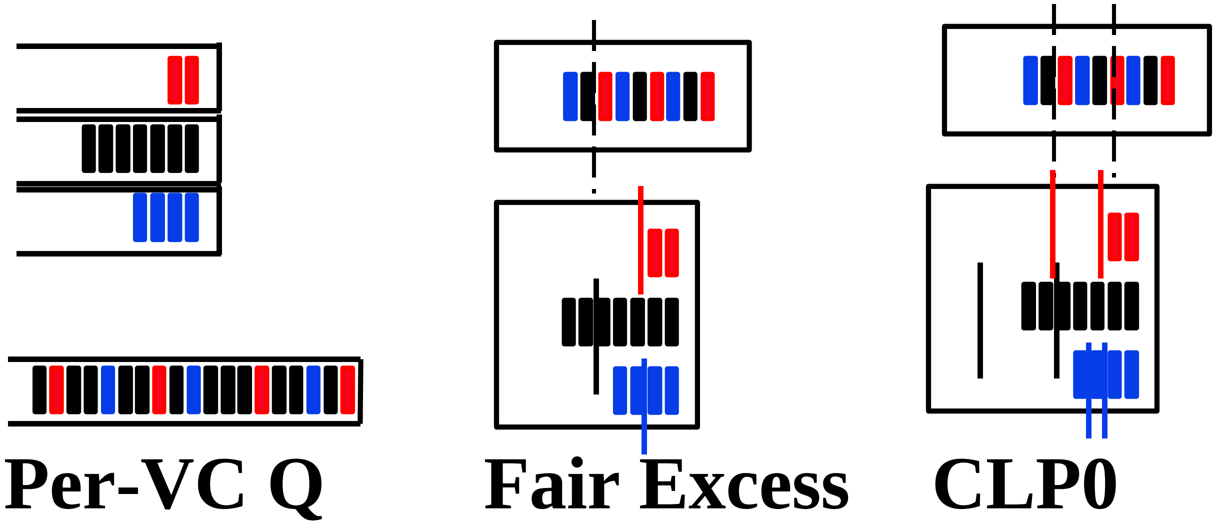


GEO: Buffers vs GR

- 5 sources, Vanilla, SACK, FRR
  SD and EPD, all average together

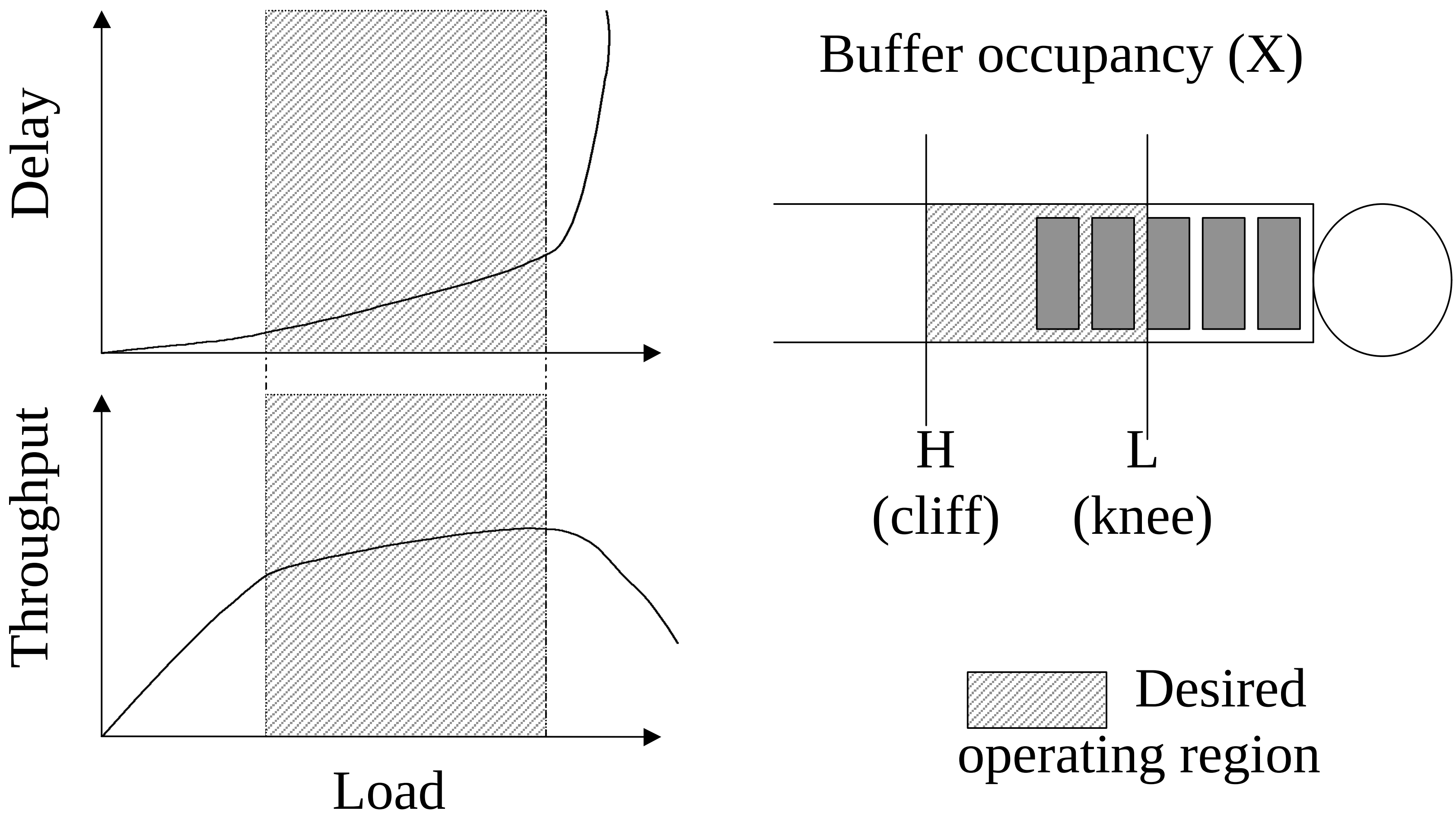


4a. GFR Results

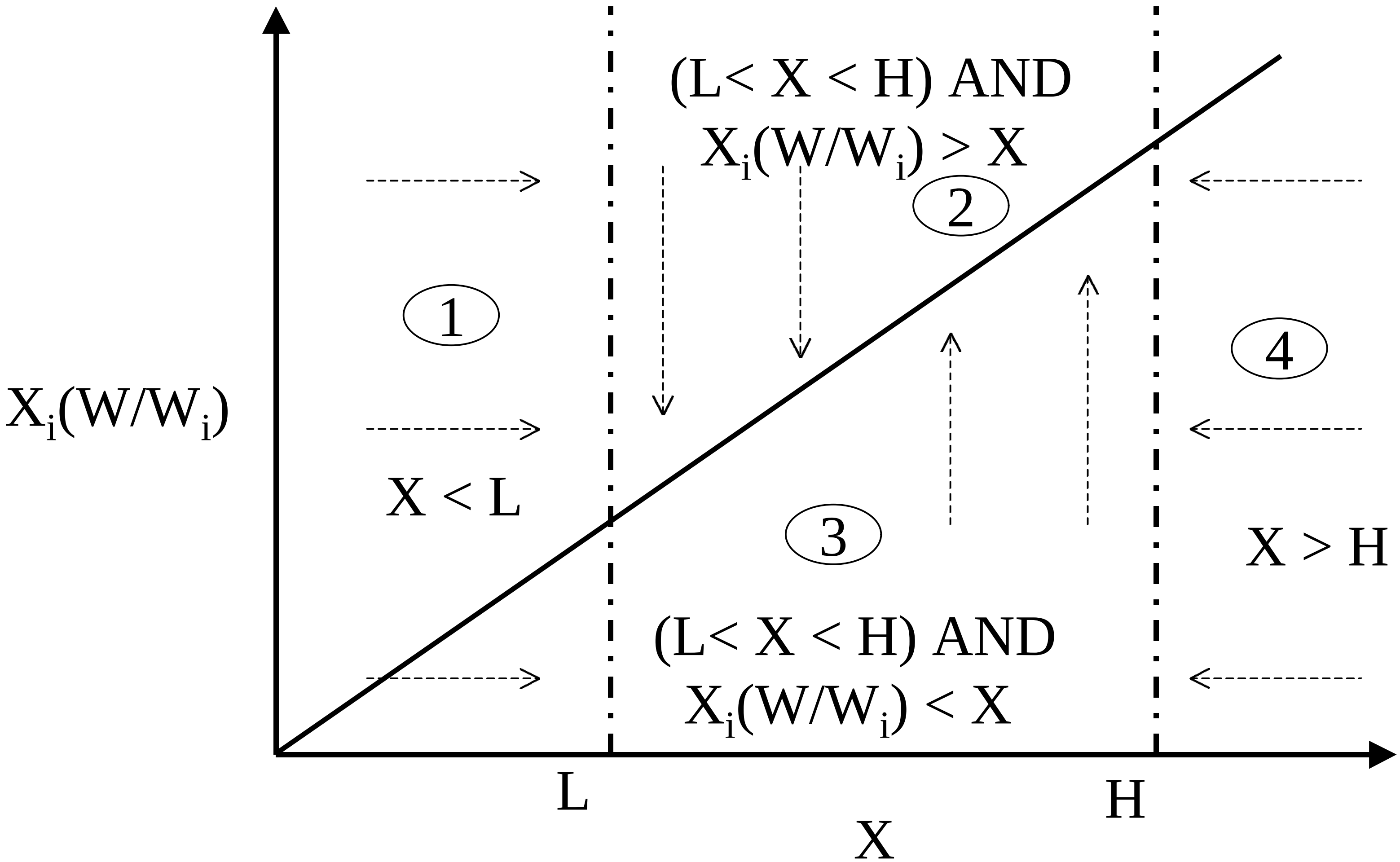


- Per-VC queuing and scheduling is sufficient for per-VC MCR.
- FBA and proper scheduling is sufficient for fair allocation of excess bandwidth
- One global threshold is sufficient for CLP0+1 guarantees Two thresholds are necessary for CLP0 guarantees

DFBA Operating Region



DFBA (contd.)



Task 6. Problem in TCP Implementations

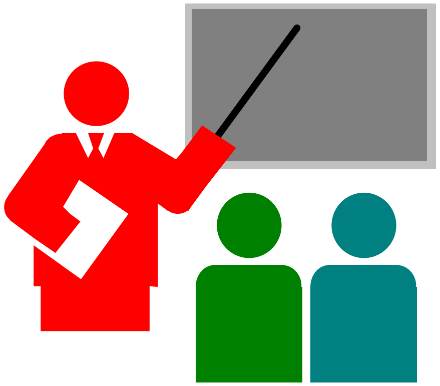
- Linear Increase in Segments:
  CWND/MSS = CWND/MSS + MSS/CWND
- In Bytes: CWND = CWND + MSS*MSS/CWND
- All computations are done in integer
- If CWND is large, MSS*MSS/CWND is zero and CWND does not change. CWND stays at 512*512 or 256 kB.

Solutions

- **Solution 1**: Increment CWND after N acks (N > 1) CWND = CWND + N*MSS*MSS/CWND
- **Solution 2**: Use larger MSS on Satellite links such that MSS*MSS > CWND. MSS $\geq$ Path MTU.
- **Solution 3**: Use floating point
- **Recommendation**: Use solution 1. It works for all MSSs.
- **To do**: Does this change TCP dynamics and adversely affect performance.
- Result: Solution 1 works. TCP dynamics is not affected.

Task 7. Optimize SACK TCP

- SACK helps only if retransmitted packets are not lost.
- Currently TCP retransmits immediately after 3 duplicate acks (Fast retransmit), and then waits RTT/2 for congestion to subside.
- Network may still be congested
  ⇒ Retransmitted packets lost.
- Proposed Solution: Delay retransmit by RTT/2, I.e., wait RTT/2 first, and then retransmit.
- Result: Delayed retransmit does not help.



Summary

Very comprehensive study of TCP/IP over UBR: existing mechanisms, new mechanisms, parameter selection

1. For satellite networks, end-system policies (SACK) have more impact than switch policies (EPD).
2. SD and FBA improve fairness.

Summary (Cont)

3. 0.5*RTT buffers provide sufficiently high efficiency (98% or higher) for SACK TCP over UBR even for a large number of TCP sources

4a: TCP throughput may be controlled with FIFO queuing. New Buffer Management Policy: DFBA.

4b. Reserving a small fraction for UBR helps it a lot in satellite networks

5. Effects of TCP mechanisms, drop policies, and buffer size for bursty traffic are similar to those for persistent TCP.

Summary (Cont)

6. Large windows cause integer division problems. Increment CWND after N acks works OK
7. Delayed retransmit has no effect.

Resulting Publications

1. Switch and end-system policies

- R. Goyal, R. Jain, S. Kota, M. Goyal, S. Fahmy, B. Vandalore, "Traffic Management for TCP/IP over Satellite-ATM Networks," To appear in IEEE Communications Magazine, March 1999, 18 pp., http://www.cis.ohio-state.edu/~jain/papers/comm399.htm
- R. Goyal, R. Jain, S. Kota, M. Goyal, S. Fahmy, B. Vandalore, "Improving the performance of TCP/IP over Satellite-ATM Networks," Under preparation. To be submitted to International Journal of Satellite Communications, Special Issue on Internet Protocols over Satellite.

Publications (Cont)

2. Switch drop policies

- R. Goyal, R. Jain, S. Kalyanaraman, S. Fahmy, B. Vandalore, S. Kota, "TCP Selective Acknowledgments and UBR Drop Policies to Improve ATM-UBR Performance over Terrestrial and Satellite Networks", Proc. ICCCN '97, Las Vegas, September 1997, pp. 17-27, http://www.cis.ohio-state.edu/~jain/papers/ic3n97.htm

3. Buffer requirements

- R. Goyal, S. Kota, R. Jain, S. Fahmy, B. Vandalore, Jerry Kallaus, "Analysis and Simulation of Delay and Buffer Requirements of Satellite-ATM Networks for TCP/IP Traffic," Under preparation, http://www.cis.ohio-state.edu/~jain/papers/jsac98.htm

Publications (Cont)

- R. Goyal, R. Jain, S. Fahmy, B. Vandalore, S. Kalyanaraman, S. Kota, P. Samudra, "UBR Buffer Requirements for TCP/IP over Satellite Networks," ATM Forum/97-0616, July 1997, http://www.cis.ohio-state.edu/~jain/atmf/a97-0616.htm

4. GFR and GR

- R. Goyal, R. Jain, S. Fahmy, B. Vandalore, "Buffer Management for the GFR Service," ATM_Forum/98-0405, July 1998, http://www.cis.ohio-state.edu/~jain/atmf/a98-0405.htm
- R. Goyal, R. Jain, S. Fahmy, B. Vandalore, "Buffer Management for the GFR Service," Submitted to Journal of Computer Communications, January 1999, 33 pp., http://www.cis.ohio-state.edu/~jain/papers/dfba_cc.htm

Publications (Cont)

- R. Goyal, R. Jain, S. Fahmy, B. Vandalore, "GFR Implementation Options," ATM_Forum/98-0406, July 1998, http://www.cis.ohio-state.edu/~jain/atmf/a98-0406.htm
- R. Goyal, R. Jain, S. Kalyanaraman, S. Fahmy, B. Vandalore, X. Cai, S. Kim, S. Kota, "Guaranteed Rate for Improving TCP Performance on UBR+ over Terrestrial and Satellite Networks," ATM Forum/97-0424, April 1997, http://www.cis.ohio-state.edu/~jain/atmf/a97-0424.htm

Publications (Cont)

5. Bursty Sources

- M. Goyal, R. Goyal, R. Jain, B. Vandalore, S. Fahmy, T. VonDeak, K. Bhasin, N. Butts, and S. Kota, "Performance Analysis of TCP Enhancements for WWW Traffic using UBR+ with Limited Buffers over Satellite Links", ATM_Forum/98-0876R1, December 1998, http://www.cis.ohio-state.edu/~jain/atmf/a98-0876.htm

6. Large congestion window

- See ICCCN'97 paper under deliverable 2 above.

7. Optimizing SACK TCP

- Analyzed delayed retransmit. No significant effect. No papers.

Thank You!